NOTICE

The invention disclosed in this document resulted from
research in aeronautical and space activities performed under
programs of the National Aeronautics and Space Administration.
The invention is owned by NASA and is, therefore, available for
licensing in accordance with the NASA Patent Licensing
Regulation (14 Code of Federal Regulations 1245.2).

To encourage commercial utilization of NASA-owned inventions,
it is NASA policy to grant licenses to commercial concerns.
Although NASA encourages nonexclusive licensing to promote
competition and achieve the widest possible utilization, NASA
will consider the granting of a limited exclusive license,
pursuant to the NASA Patent Licensing Regulations, when such a
license will provide the necessary incentive to the licensee to
achieve early practical application of the invention.

Address inquiries and all applications for license for this
invention to NASA Resident Office-JPL, NASA Patent Counsel,
Mail Code 180-801, 4800 Oak Grove Dr., Pasadena, CA 91103.
Approved NASA forms for application for nonexclusive or
exclusive license are available from the above address.

Inventors:
   Alexander W. Moopenn
   Anilkumar P. Thakoor
   and John J. Lambe

JPL Case No. 17058
NASA Case No. NP-17058-1-CU
May 21, 1987
S. N. 060,201
Filing Date: 6-10-87

Contractor: Jet Propulsion Laboratory

HYBRID ANALOG-DIGITAL ASSOCIATIVE NEURAL NETWORK

The instant invention is directed to hardware for investigation of the dynamics, associative recall properties and computational capability of neural networks.

The Hopfield model of a neural network memory shown in FIG. 1 is implemented as shown in FIG. 2 with random access memory (RAM) that can be expanded, such as one N x M bit RAM 1 expanded by a second N x M bit RAM 2 to store a matrix of N x N bits (which represent synapses) and an analog circuit comprised of switches SW_1-SW_N and resistors R_1-R_N connected to a summing amplifier 4. As each row of synaptic information is read out of the RAM, signals stored in an array 8 of sample-and-hold devices S/H2 are coupled by the resistors R_1-R_N and summed in an amplifier 4. The sum is sampled and stored in successive ones of an array 5 of sample-and-hold devices S/H1 implemented as shown in FIG. 2 with an RC circuit for smoothing. This successive sampling of row sums is controlled by a shift register 6,7 which shifts a sample control bit 1 in response to a timing signal ϕ_1. The sums of all rows are processed by neuron amplifiers A_1-A_N and sampled in parallel by an array 8 of sample-and-hold devices S/H2 in response to a timing signal ϕ_2 (see FIG. 2b). The process is repeated until the output of the neuron amplifiers stabilize. An INHIBIT signal sets a threshold level for the neuron amplifiers. Initial PROMPT signals are entered through the neuron amplifiers. To allow for negative synaptic values, two RAMs 10 and 11 are used. The RAM 10 is used for synaptic values of 0 and +1, and the RAM 11 is used for synaptic values of -1. Thus, for a synapse of -1, a bit 1 is stored in the RAM 11; otherwise a bit 0 is stored. For a synapse of +1, a bit +1 is stored; a bit 0 is otherwise stored. Separate summing amplifiers 12 and 13 are provided for the positive and negative synaptic matrices. The outputs of the two summing amplifiers are then combined in a difference amplifier 14 which corresponds to amplifier 4 in FIG. 2. The systems of FIG. 2 and FIG. 3 are otherwise the same. FIG. 3 illustrates how control may be provided by a microprocessor, μP. The same microprocessor control could be applied to the system of FIG. 2.

The novelty of the invention resides in storing synaptic information in random access memory, and in using analog circuits for effecting feedback to the neuron amplifiers, but with digital control over sample-and-hold circuits so that the process may be stopped at any cycle for analysis. The synaptic information may be easily changed to present a new problem, or to vary a problem. The problem may be modeled for investigation or may be one for solution.

FIG. I (PRIOR ART)

FIG. 2b

FIG. 2

FIG. 2a
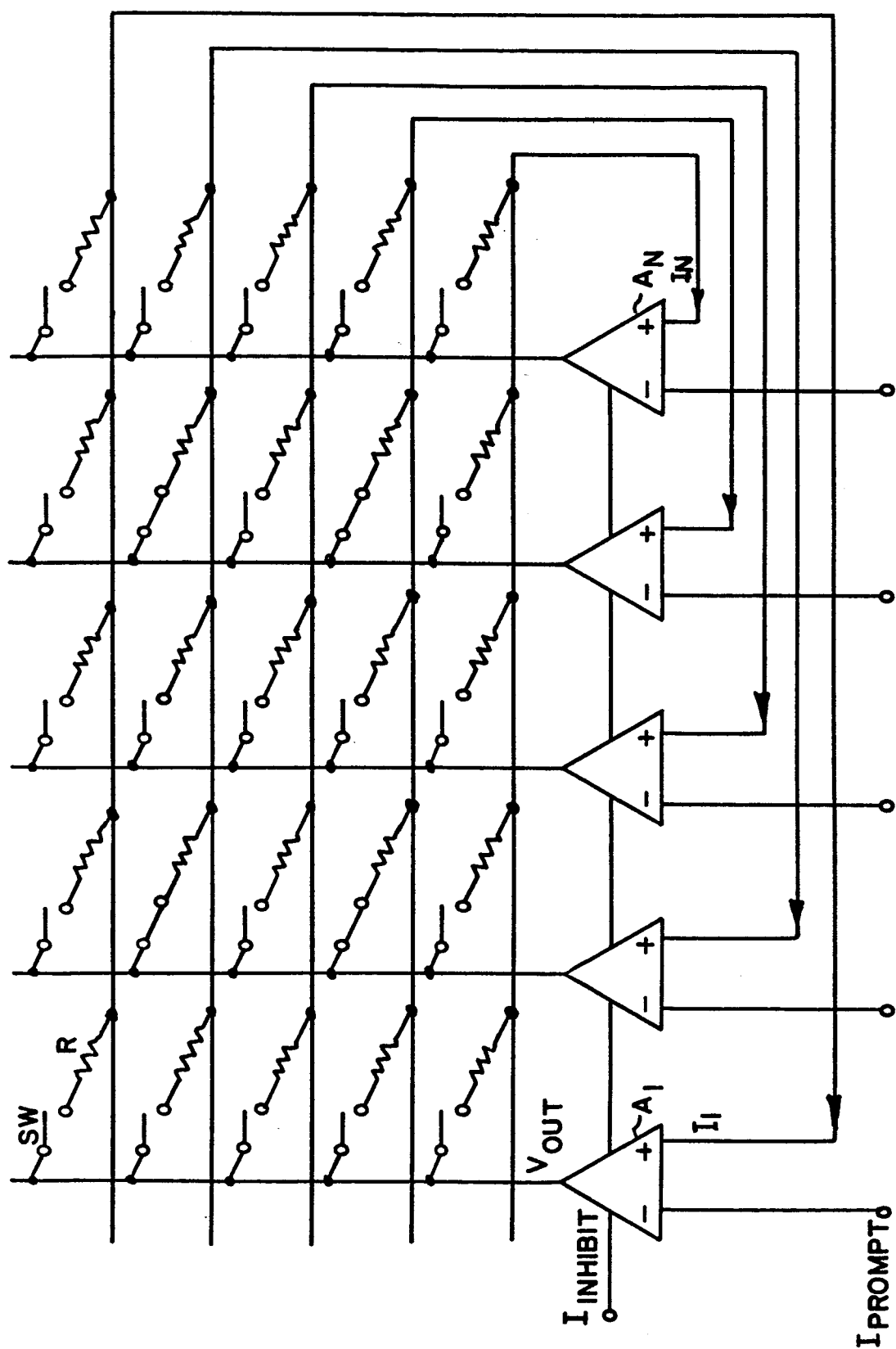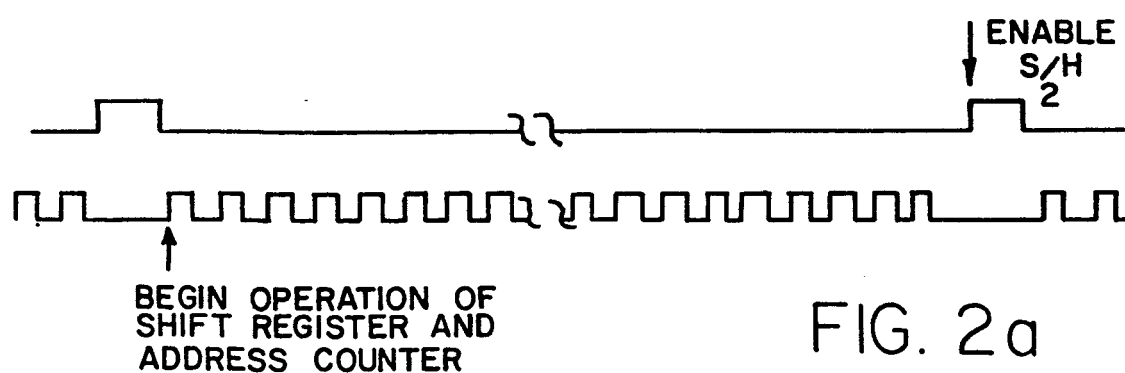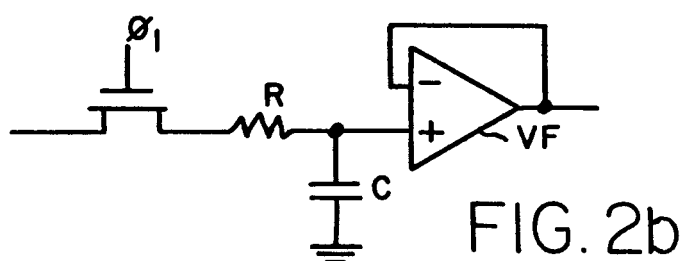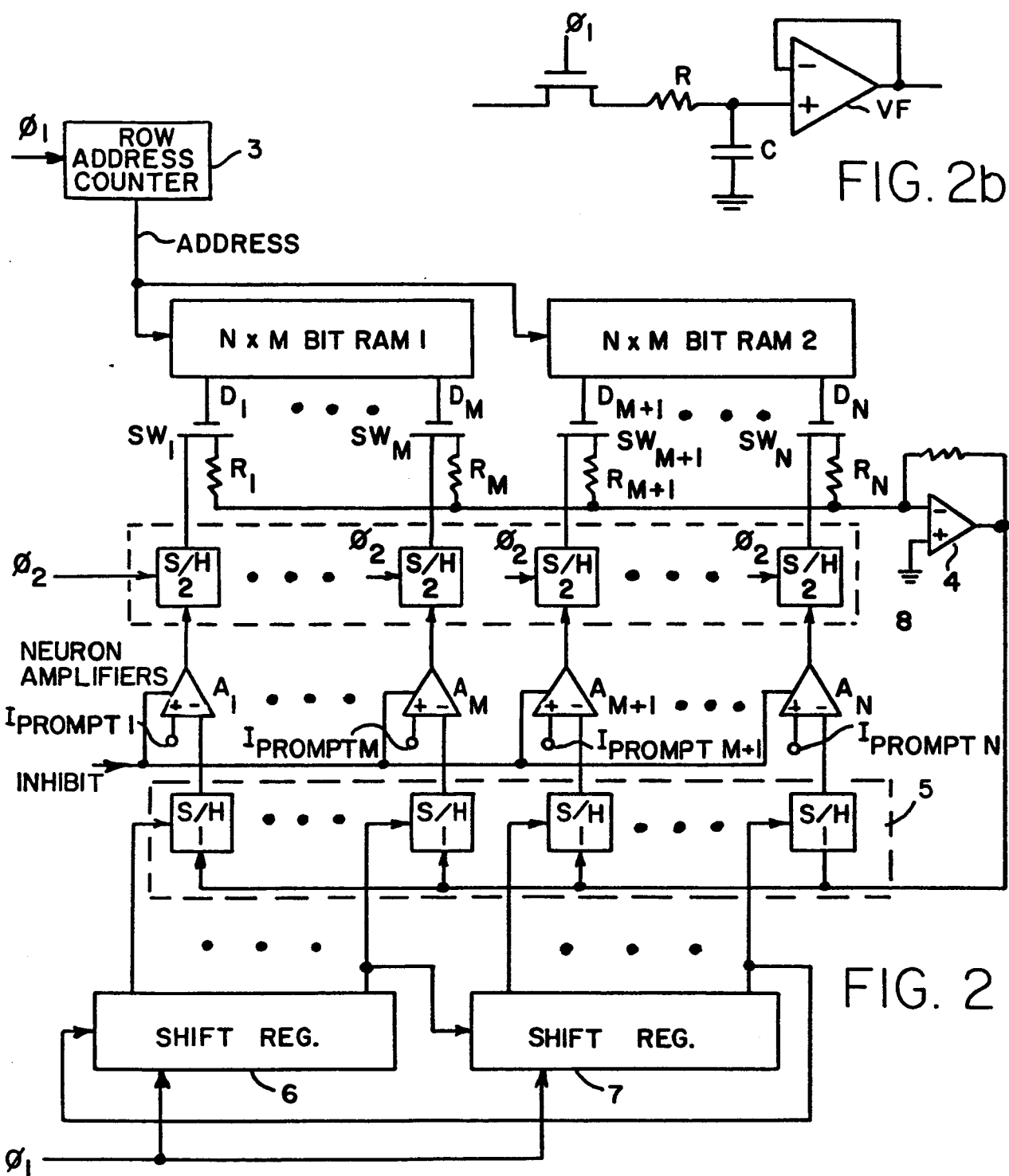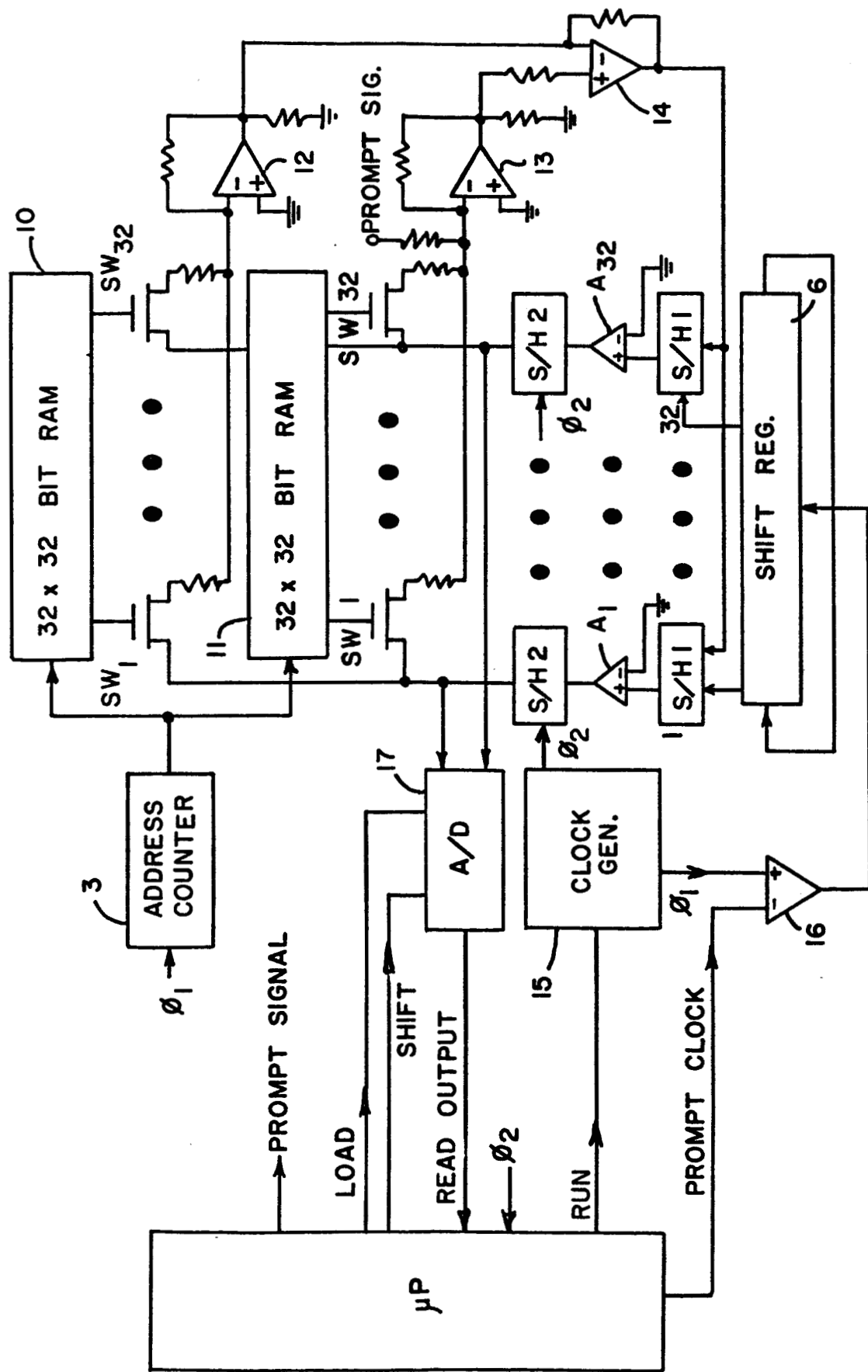
FIG. 3

HYBRID ANALOG-DIGITAL ASSOCIATIVE NEURAL NETWORK

ORIGIN OF INVENTION

The invention described herein was made in the performance of work under a NASA contract, and is -subject to the provisions of Public Law 96-517 (35 USC 202) in which the Contractor has elected not to

5   retain title.

BACKGROUND OF THE INVENTION

This invention relates to an analog-digital hybrid neural network realization of associative memories and collective computation systems based on the Hopfield model disclosed in U.S. Patent No.

10 4,660,166.

The Hopfield model shown in **FIG. 1** in its simplest form defines the behavior of a state vector V as a result of the synaptic interactions between its components $I_1 - I_n$ using a matrix of row and

15 column conductors, where the columns (or **rows**) are driven by neuron amplifiers A_1-A_n with feedback from rows (or columns) through synapses at intersections between the rows and columns. Connections are selectively made through resistors **R** by switches **SW** to

20 store synaptic information in the matrix. The value of each resistor **R** is preferably selected for the problem at hand, but may be made the same for every synaptic node thus created. Each amplifier drives the synaptic nodes (resistors) to which the amplifier

25 output is connected by column conductors, and receives feedback from every synaptic node to which a row conductor is connected. Each amplifier thus acts as a neuron capable of feeding back to all other

neurons through synaptic nodes connected to the neuron output by a resistor **R**. The final output of the neuron amplifiers is a set of voltage signals that represents a stored word which best matches an input

5 word consisting of a set of signals $I_{PROMPT\ 1}$ through $I_{PROMPT\ N}$.

A neural network having a 32 x 32 matrix of synapses has been implemented for research purposes with electronic switches and resistors of equal value

10 for all synaptic nodes, and using analog amplifiers with an inhibit input for control of its threshold. It would be desirable to expand the network to a much larger matrix, such as 1000 x 1000, but such a large network would require too many switches and resis-

15 tors.

In the continuing research effort into associative memories and collective computation systems based on the Hopfield model, the need has arisen for a research network utilizing hardware that can offer

20 higher operating speeds than is presently obtainable through computer simulation, and that can be easily scaled upwards, e.g., scaled up from 32 to 1000 or more neurons. Software simulations, while proving a concept and verifying the basic operating principles

25 of the Hopfield model, suffer from intrinsic speed limitations as compared to the expected speed (cycle time) of neural networks embodied in hardware. Thus, in order to research the properties and requirements of neural network associative memories and computers,

30 programmable, flexible, and easily expandable neural networks using state-of-the-art techniques and components are required.

SUMMARY OF THE INVENTION

In accordance with the present invention, an analog-digital hybrid network utilizes a block of NxN random access memory (RAM) for storing a matrix of binary synaptic information, a linear array of N

5 analog amplifiers for performing the neuron function, an array of N electronic switches, one switch for each column of the synaptic information matrix for connecting the neuron amplifier feedback signal to a single row conductor, which connects to all of the

10 neuron amplifiers through a summing amplifier, and two arrays of N sample-and-hold means, a first array for integrating, sampling and holding the output signal of the summing amplifier for each row as an input to each of the neuron amplifiers, and a second

15 array of nonintegrating sample-and-hold devices for holding the respective neuron amplifier output signals for a complete matrix cycle to the synaptic switches. A microcomputer addresses the ROM to present to the synaptic switches one row of synaptic data

20 at a time from the matrix. Thus, the feedback from successive rows of synaptic information produced is distributed to the first array of integrating sample-and-hold devices sequentially where they are stored as inputs to the neuron amplifiers until all rows of

25 synaptic information in the RAM have been stored. Then the outputs of the neuron amplifiers are stored in the second array of sample-and-hold devices. The first array of sample-and-hold devices thus smooths the feedback for each column signal from the summing

30 amplifier of all rows, and the second array holds the feedback signals from all rows of synaptic information. The microcomputer may read out the signals from the second, nonintegrating array of sample-and-

hold devices as the output of the neural network once
all of the rows of synaptic information stored in the
RAM have been processed through a sufficient number
of matrix cycles for the output to stabilize.   A
5   second RAM, a second row of synaptic switches, and a
second summing amplifier may be provided, and the
difference between the two amplifiers is generated by
a differential amplifier.   That difference is then
used as the feedback to the neuron amplifiers.   In
10 that way each bit of synaptic information need not be
limited to two levels, 1 or 0, but may be trilevel,
1, 0 or -1. This synaptic network may easily be ex-
panded from, e.g., 32 x 32 to 1000 x 1000 by merely
substituting larger RAMS and extending the synaptic
15 hardware linearly from 32 to 1000 synaptic switches,
neuron amplifiers, and sample-and-hold devices for
each column of synaptic information.

     The novel features that are considered charac-
teristic of this invention are set forth with par-
20 ticularity in the appended claims.   The invention
will best be understood from the following descrip-
tion when read in connection with the accompanying
drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

     FIG. 1 is a schematic diagram of a synaptic
25 network with a matrix of switches and resistors for
selecting the synaptic nodes, and five neuron ampli-
fiers, in accordance with the prior art.

     FIG. 2 is a schematic diagram of a neural
network computer embodying the architecture of the
30 present invention from binary synaptic data bits of 1
and 0, **FIG. 2a** is a timing diagram for the operation
of the network of **FIG. 2**, and **FIG. 2b** is a schematic

circuit diagram for integrating sample-and-hold (S/H) devices used in a first array of S/H devices.

FIG. 3 is a schematic diagram of a neural network computer embodying the architecture of the

5   present invention for trilevel synaptic information bits of 1, 0 and -1.

DESCRIPTION OF PREFERRED EMBODIMENTS

Referring first to the prior art **FIG. 1**, it may be appreciated that as the size of the network is expanded to N x N, where N may be 1000 or larger, the

10   number of switches required to store the synaptic data expands as the square of N, while the number of neurons increases linearly as a function of N. Once the input vector is applied in the form of N synaptic data signals, I_{PROMPT}, the network will seek a solu-

15   tion in the form of an output vector appearing on vertical conductors where the signal on **each vertical** conductor is a stable maximum or minimum. The neuron amplifiers may have an inhibit input to control the threshold of the amplifier so that its output will

20   remain a steady minimum, until the input to the amplifier exceeds a threshold level. This will aid the network in reaching a stable state more quickly. While such an analog neural network will reach a solution in a minimum of time, it does not lend it-

25   self well to expansion for research purposes.

The ability to expand a neural network is achieved by an analog-digital hybrid architecture shown in **FIG. 2** using a random access memory to store a matrix (N x N) of synaptic data bits having values

30   of 1 and 0. To illustrate the ease with which the network may be expanded, it is illustrated as having already been expanded from M x M to N x N, where N =

2M, by adding to a first RAM 1 a second RAM 2 where
each RAM is N x M. A row address counter 3 responds
to clock signals ϕ_1 to advance the address of the
memories row by row through N rows of synaptic infor-

5   mation. The outputs of the two RAMs labeled D_1
through D_N for each row control a set of N pass tran-
sistor switches SW_1-SW_N to pass current through re-
spective resistors R_1-R_N when the controlling
synaptic bit read is equal to one and to block cur-

10 rent when the synaptic bit is equal to zero at each
position D_1 through D_N.

The synaptic resistors are connected to a
summing amplifier 4 where they are added. The sum is
then sampled and held for each row of the RAM by an

15 array 5 of integrating sample-and-hold devices S/H1,
selected one at a time for each row of synaptic in-
formation processing switches. To accomplish that, a
serial shift register advances a bit 1 in a pair of
cascaded M-bit shift registers 6 and 7 in response to

20 a clock signal ϕ_1, where $\phi_1 = N\phi_2$, as shown in the
timing of **FIG. 2a.** In that way the synaptic matrix
sum for every row of the matrix is stored, one sum at
a time, in successive ones of a linear array of sam-
ple-and-hold devices S/H1.

25 The address counter 3 is advanced by one to
address the next row of synaptic information stored
in the memory each time a sum from the amplifier is
stored in a sample-and-hold device. At the same
time, the clock signal ϕ_1 shifts the bit 1 in the

30 shift register to the next position, until the bit 1
has been shifted to the Nth position, and the Nth row
of synaptic information has been read from memory to
complete a matrix cycle. At that time, each sample-
and-hold device in a position j in the array 5 holds

a signal that is a function of the jth row of infor-
mation. After a clock signal ϕ_1 has been generated N
times, a nonoverlapping ϕ_2 pulse is generated, as
shown in **FIG. 2a**, which completes a matrix cycle.

5 Another cycle may then follow.

The signals from the first array **5** of sample-
and-hold devices are passed by an array of neuron
amplifiers **A$_1$** through **A$_N$**, each having an **INHIBIT**
input signal applied which sets their threshold

10 value. The signals which are passed by the neuron
amplifiers are then sampled and held by a second
array **8** of nonintegrating sample-and-hold devices
S/H2 at the time of initiating the next matrix cycle
of synaptic information processing. The process is

15 repeated for the next and each subsequent row, each
time storing the results in the array of integrating
sample-and-hold devices **S/H1**, and then transferring
those results to the second array of nonintegrating
sample-and-hold devices **S/H2** at the start of the

20 procedure for the next matrix cycle. **FIG. 2b** illus-
trates schematically the implementation of an inte-
grating sample-and-hold circuit with a switch **SW, RC**
circuit, and noninverting voltage follower **VF**. The
entire process for all rows in the stored matrix of

25 synaptic data is repeated, each time updating the
signals stored in the nonintegrating sample-and-hold
devices **S/H2**, until the output of the summing ampli-
fier **4** becomes stable from one matrix cycle to the
next of the entire process.

30 To enter a problem solving operation, an ini-
tial set of signals **I$_{PROMPT\ 1}$** through **I$_{PROMPT\ N}$** is
entered through the neuron amplifiers **A$_1$** through **A$_N$**,
and stored in the array of nonintegrating sample-and-
hold devices **S/H2**. This is accomplished by generat-

ing a signal ϕ_2 which enables all the sample and hold
devices **S/H2.** Once the "prompt" is entered, it is
removed from the neuron amplifiers. In terms of an
associative memory, the "prompt" may be "clues" of
5   information stored in the memory. The output of the
integrating sample-and-hold devices **S/H2,** after sys-
tem stability is reached, will be the best match to
⁻the "prompt" from all the data stored.

In summary, system operation is controlled by
10   two clock signals, ϕ_1 and ϕ_2, with ϕ_1 = $N\phi_2$, where N
denotes the number of neurons N = 2M.   On each ϕ_1
timing signal, a neuron amplifier is selected by the
shift registers **6, 7** and its input is updated and
sampled by integrating sample-and-hold devices **S/H1.**
15   During the update process, the neuron input signal is
determined as a function of a row of synaptic infor-
mation which controls the transistor switches D_1-D_N
and the present neuron output state at the outputs of
the nonintegrating sample-and-hold devices **S/H2.** This
20   process is repeated for each row of synaptic informa-
tion in the memory following each ϕ_1 timing signal
updating the sample-and-hold device **S/H1** in turn as
directed by the serial shift registers until all rows
of the synaptic information have been processed. On
25   each ϕ_2 clock pulse, the neuron amplifier outputs
(determined by their updated inputs and the inhibit
signal) are sampled by the array of nonintegrating
sample-and-hold devices **S/H2,** completing one matrix
cycle.   Thus the system cycle time is determined by
30   the update ϕ_1 timing signal rate, and the number of
neurons.

An advantage of this architecture is that the
physical interconnections of the synaptic matrix of
FIG. 1 are not required, thus vastly reducing system

complexity, which permits very large neural networks to be simulated. Although its use is contemplated for research, future developments in the use of the present invention can conceivably follow with minor modi-

5   fications of the hybrid architecture disclosed. One such modification, already designed, fabricated and tested, is an expansion of the concept to accommodate three quantized levels of +1, 0 and -1, as shown in **FIG. 3.**

10   Referring to **FIG. 3,** a 32 x 32 memory **10** is used in the same way as the RAMs **1** and **2** of **FIG. 2.** That memory could be made up of two 32 x 16 memories, or could be expanded to a 64 x 64 memory with two 64 x 32 RAMs.   However, in practice, a sufficiently

15   large single RAM is selected, such as 256 x 256 or 1024 x 1024. This memory **10** stores the synaptic bits of value **0** and **1**. A second memory **11** stores synaptic bits of value **0** and **1**.

           For each bit position of the memory **10** there

20   is stored a bit value of **0** or **1**, and for each bit **1** stored in memory **10,** there must be stored at the same address of memory **11,** a bit **0**.   Similarly, for each bit position of the memory **11,** there is stored a bit value of **0** or **1**, and for each bit **1** stored in memory

25   **11,** there must be stored at the same address of memory **10** a bit **0**.   Summing amplifiers **12** and **13** respond in the same manner as the amplifier **4** in **FIG. 2,** but their outputs are connected to separate terminals of a difference amplifier **14**.   The summing amplifiers

30   are inverting operational amplifiers, so the amplifier **12** which sums the synaptic results of a bit **1** from the memory **10** is connected to the inverting (-) input terminals of the difference amplifier **14,** and the amplifier **13** which sums the synaptic results of a

bit **1** from the memory **11** is connected to the nonin-
verting (+) input terminal of the difference ampli-
fier **14**. The architecture of the trivalue embodiment
is otherwise the same as illustrated in **FIG 2**. The
5   components which remain the same are identified by
the same reference characters. But note that the
inhibit function, i.e., the setting of the threshold
value for the neurons (amplifiers A_1-A_N) is now ef-
fectively obviated by allowing for negative values of
10 synaptic information bits, and using the summing
amplifier **13** to subtract the results of negative
synaptic values from the results of positive synaptic
values, thereby transferring the inhibit function to
a point ahead of the sample-and-hold devices **S/H1**.

15   In summary, the architecture of **FIG. 3** util-
izes two memory blocks, one for implementing the
excitation synapse function (response to synaptic
data bits of value +1) and the other for implementing
the inhibition synaptic functions (response to synap-
20 tic data bits of value -1).   This allows a hybrid
implementation of the Hopfield model disclosed in the
U.S. patent 4,660,166 with each synaptic data bit of
strength (T_{ij}) having a value of +1, 0 or -1. These
synaptic strength values are readily programmed in
25 the memories **10** and **11** to provide any synaptic matrix
desired. Additional memory blocks, switches connect-
ing resistors and summing amplifiers could be pro-
vided to increase the number of grey levels in the
synaptic strength of each matrix (memory) location.

30   In operation, a neuron amplifier A_1 is select-
ed by the output of the serial shift register **6**, and
the input of the neuron amplifier is sampled and held
by its integrating sample-and-hold device **S/H1**. The
input signal level during its update cycle is deter-

mined by the analog output states of the neuron am-
plifiers A_1-A_{32} stored in nonintegrating sample-
and-hold devices **S/H2** during the previous ϕ_2 clock
signal, and the row of synaptic information appearing

5   at the memory outputs which control the switches
SW_1-SW_{32} for the resistor connections. On each ϕ_2 cy-
cle, the neuron outputs, as determined by their up-
dated input signal and threshold, are sampled and
held by their output sample-and-hold devices **S/H2**,

10   thus completing one iterative or system cycle. This
process is repeated until all neuron amplifier out-
puts have been stabilized.

The hybrid architecture is ideally suited for
simulating both continuous and discrete neural net-

15   works of the Hopfield model. In the continuous net-
work, the input voltage of the ith neuron on the kth
iterative cycle, $u_i(k)$, is given by

$$u_i(k) = u_i(k-1) + [\Delta t/\tau_i) \sum_j T_{ij}V_j(k-1) - u_1(k-1)]$$

where τ_i is the time constant of the neuron (ampli-

20   fier i) as determined by the RC product of the neu-
ron's input sample/hold device, V_j is the analog
output of the jth neuron (amplifier j) and Δt is the
input sampling time. In the discrete model, $u_i(k)$ is
given by

25          $$u_i(k) = \sum_j T_{ij}V_j(k-1)$$

The neuron output, $V_i(k)$, in all cases, is obtained
from

$$V_i(k) = f_i[u_i(k)-u_{io}]$$

where u_{io} and f_i are, respectively, the threshold and the transfer function of the comparator for the ith neuron.

5   The system cycle time is basically given by $T = N\phi_1$. Thus, for a given N, the operation speed of the hybrid architecture is primarily governed by the sampling times. Moreover, in order to simulate a neural network accurately in the continuous model,
10   the sampling time should be small compared to the time constant of the neurons. In the prototype system of **FIG. 3**, a conservative input sampling time of about 8 microseconds and a neuron time constant of one millisecond have been chosen. However, the ar-
15   chitecture can be further optimized with custom VLSI chips to reduce the input sampling time to a microsecond so that for N=1000, a system cycle time on the order of a millisecond can be achieved.

       In the implementation of this embodiment, a
20   programming microprocessor µP controls operation by a signal **RUN** which enables a clock generator **15** to transmit ϕ_1 timing signals to the shift register **6** via an OR gate **16** and the address counter **3**, and to transmit ϕ_2 timing signals to the sample-and-hold
25   devices **S/H2**, and optionally to the microprocessor, which may be programmed to initiate another **RUN** signal that starts another clock generator cycle or, in the case of a free running clock generator, terminating the **RUN** signal to stop the clock generator until
30   it is determined that the system should continue to run. In either case, the microprocessor may control a sample-and-hold A/D converter **17** having a multiplexer input and a serial shift register output, to

convert the output of the sample-and-hold devices S/H2 in sequence from analog to digital form in order that the output be read into the microprocessor µP.

The microprocessor may then determine whether
5   stability has been reached. To allow for A/D conversion, reading the output, and time to compare the output read with an output previously read, the microprocessor may be programmed to hold off issuing another **RUN** signal, or in the case of a free running
10   clock generator, interrupting the **RUN** signal while the comparison is made. Once it is determined that the system output has stabilized, the last output read is the information desired. It would also be possible to monitor the output of the amplifier **4** in
15   **FIG. 2** and the amplifier **14** in **FIG. 3** to determine when stability has been reached. This option would be used for visual monitoring through a volt meter connected to the amplifier output.

To enter a "prompt" by the programmed micro-
20   processor, a **PROMPT CLOCK** is generated by the microprocessor before any **RUN** signal is generated. At the same time, a **PROMPT SIGNAL** is generated as a train of pulses applied through the amplifier **13** and amplifier **14** to the sample-and-hold devices **S/H1**. The clock
25   generator **15** is then designed to generate an initial ϕ_2 timing signal which transfers the distributed "prompt" signals into the second array of sample-and-hold devices **S/H2**.

In addition to the ease of expandability as
30   compared to a fully parallel neural network hardware and improved speed with respect to most of the current software simulations, the hybrid neurocomputer provides extreme versatility in its configurability. Furthermore, like the software simulations, the op-

eration of the hybrid neurocomputer can be "frozen" at any stage or at each iteration cycle to study the dynamic behavior of the neurons as the circuit updates.

5   The hybrid architecture is proving to be a very versatile and powerful neural network simulator in several different configurations for research. The performance of the 32-neurons prototype hybrid system has been tested for a variety of applications includ-

10   ing associative memory with global and local inhibitions, destabilization of spurious states using asymmetric connections, and combinatorial optimization problems such as Map Coloring.

      Although particular embodiments of the inven-

15   tion have been described and illustrated herein, it is recognized that modifications and variations may readily occur to those skilled in the art. Consequently, it is intended that the claims be interpreted to cover such modifications and variations.

15
20

ABSTRACT OF THE DISCLOSURE

HYBRID ANALOG-DIGITAL ASSOCIATIVE NEURAL NETWORK

Random access memory is used to store synaptic
information in the form of a matrix of rows and col-
umns of binary digits.   N rows read in sequence are
processed through switches and resistors, and a sum-
5   ming amplifier to N neural amplifiers in sequence,
one row for each amplifier, using a first array of
sample-and-hold devices S/H1 for commutation.   The
outputs of the neural amplifiers are stored in a
second array of sample-and-hold devices S/H2 so that
10   after N rows are processed, all of said second array
of sample-and-hold devices are updated.   A second
memory may be added for binary values of 0 and -1,
and processed simultaneously with the first to pro-
vide for values of 1, 0, and -1, the results of which
15   are combined in a difference amplifier.